REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy)	2. REPORT TYPE	3. DATES COVERED (from. . . to)
08/08/01	FINAL	2/12/01 – 8/12/01

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Mission Readiness Measurement Aid (MIRMAID)	N00024-01-C-4040
	5b. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5c. PROJECT NUMBER
Tim Bowden and Ed Brumbaugh	5d. TASK NUMBER
	5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Micro Analysis and Design, Inc. 4900 Pearl East Circle Suite 201E Boulder. CO 80301	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. MONITOR ACRONYM
Naval Sea Systems Command Mr. Barry Stevens 1922 Regulius Ave., Bldg 127-S / S281 Virginia Beach, VA 23451-2097	11. MONITOR REPORT NUMBER

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*:

The purpose of MIRMAID is to evaluate the overall readiness to perform a specific mission aboard any Naval vessel. The tool we have designed is intended to combine automated and observed measures of performance to provide the Commanding Officer feedback regarding the readiness of his unit to perform key missions as defined by the Uniform Naval Task List (UNTL). A primary strength of this tool is that it traces tasks identified in the UNTL down through mission essential tasks to specific operator actions and then links these operator actions to triggers in the combat system, which can initiate the data collection process. In addition to this hierarchical decomposition of UNTL tasks, MIRMAID uses scenario data output from BFTT to compare actions and decisions made by the operator to the ground truth of the scenario. A final strength of the software is that through human performance modeling, MIRMAID will be capable of comparing actual operator performance with that of an "expert" and provide both the responsible evaluator and the operator with performance feedback identifying areas where performance can be improved.

15. SUBJECT TERMS

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unclassified	25	Tim Bowden, (407) 482-6404

Ship Mission Readiness Measurement System

The Mission Readiness Measurement Aid (MIRMAID) SBIR Topic N00-123

Phase I Final Report

Prepared for:

Naval Sea Systems Command

Prepared by:

Micro Analysis and Design, Inc.
Basic Commerce and Industries, Inc.

September 12, 2001

Micro Analysis & Design

Table of Contents

Table of Figures

List of Tables

Ship Mission Readiness Measurement System (MIRMAID)
SBIR Topic N00-123
Phase II Proposal Outline

1. Project Summary

The purpose of MIRMAID is to evaluate the overall readiness to perform a specific mission aboard any Naval vessel. The tool we have designed is intended to combine automated and observed measures of performance to provide the Commanding Officer feedback regarding the readiness of his unit to perform key missions as defined by the Uniform Naval Task List (UNTL). A primary strength of this tool is that it traces tasks identified in the UNTL down through mission essential tasks to specific operator actions and then links these operator actions to triggers in the combat system, which can initiate the data collection process. In addition to this hierarchical decomposition of UNTL tasks, MIRMAID uses scenario data output from BFTT to compare actions and decisions made by the operator to the ground truth of the scenario. A final strength of the software is that through human performance modeling, MIRMAID will be capable of comparing actual operator performance with that of an "expert" and provide both the responsible evaluator and the operator with performance feedback identifying areas where performance can be improved.

MIRMAID will act both as a decision aid for commanding officers in evaluating the crew's readiness to perform specific mission requirements, and as an aid to fleet training staff in identifying specific areas where a particular crew needs improvement. Commanding officers and training personnel will possess a valuable tool for guiding the expenditure of limited training and combat resources with the deployment of the readiness measurement aid.

2. *Overview*

This document has two main purposes:

1. To describe the findings of Phase I of the project, and
2. To provide a functional/conceptual design for the tool we propose to build and deliver. This design and the associated report are the primary deliverables of the Phase I effort.

In this section we provide a brief discussion of the history of the project and identify the need for the MIRMAID tool. We also discuss the potential user population for MIRMAID.

The remaining sections are:

- Section 3 – Describes the work performed during the Phase I contract. This section also includes a description of the tasks that were accomplished and the products of these tasks.
- Section 4 – Contains the primary deliverable for the Phase I contract; the MIRMAID functional description.
- Section 5 – Contains the final monthly progress report for the Phase I project.

2.1 History of the project

In March 2001 Micro Analysis and Design (MA&D) and Basic Commerce and Industries (BCI) were awarded a Phase I Small Business Innovative Research (SBIR) contract entitled "Ship Mission Readiness Measurement System (MIRMAID)." The purpose of the project was to study the feasibility of and create a conceptual design for an integrated performance measurement system to support commanding officer evaluations of crew readiness. This tool can be used both in training exercises to identify specific areas for improvement, e.g., direct training resources, and when deployed to assess the probability of success for a given mission. MIRMAID will provide a needed advancement in readiness assessment through adding significant automated assessment and feedback to the process.

Over the past ten years our military forces have been stretched extremely thin, as global hotspots have demanded the attention of U.S. Forces. This worldwide deployment coupled with an emphasis on decreased defense spending and reduced manning has made maintaining readiness a formidable challenge. Add to this the increasing difficulty of recruiting and retaining personnel and you have an equation that may add up to unfavorable conditions of readiness. As automation on Navy ships increases, assessing the readiness of crew to operate these new, complex automated systems becomes more and more difficult. Often the assumption made is that the new automation will increase a crew's ability to perform a given mission. We know, however, from research in aviation and process control that increasing automation often compromises system safety rather than enhancing it as intended (Woods, Potter, Johannesen & Holloway, 1991; Wiener, 1985).

The military environment today is characterized by change and uncertainty. These circumstances may create conditions of reduced readiness aboard Navy vessels and in

other military installations. The need clearly exists for commanding officers to have access to concrete and specific information regarding the readiness of their personnel to accomplish a given mission. There are currently methods in place, which would allow the collection of this much-needed information. However, these systems tend to be labor intensive, time consuming and by in large subjective. What is needed is an integrated tool set that provides objective measures of crew readiness in an efficient and effective manner. The aim of the MIRMAID tool is to provide a quick, easy look at crew readiness for accomplishing specific mission essential tasks. The tool will provide the Commanding Officer with the ability to make well-informed decisions based on specific information about the readiness of his personnel.

2.2 Related Efforts

MA&D has performed the following projects which provide us with an excellent foundation on which to build the MIRMAID system.

2.2.1 Shipboard Instructor Training and Support (SITS)

The Navy has been moving towards the use of embedded and on-board training systems as a means of providing ongoing shipboard training to maintain combat readiness. However, to be used effectively and efficiently, on-board training systems require tools and methods to support and train the shipboard instructors who will use these systems. The Shipboard Instructor Training and Support (SITS) program was part of an ongoing effort to employ current research and technology in the development of solutions to continuously improve the efficiency and effectiveness of Navy shipboard training. Micro Analysis & Design was a sub-contractor on this project.

The SITS program objectives involved both instructor support and instructor training. In the area of instructor support, a SITS program objective was to investigate methods and strategies for improving team performance through the use of training technologies that support the instructor's role of guiding learning. Previous research suggested that the areas where technological support could be provided include: automated performance recording and outcome assessment, diagnosis of performance process and identification of training needs, preparation and delivery of feedback, and implementation of instructional strategies through scenario design and modification.

2.2.2 Composable Behaviors in Joint Semi-Automated Forces

This project is currently in the development phase. It will provide the capability to link entity behavior in the synthetic environment of Distributed Simulations with variable fidelity in human performance and system models. MA&D's approach includes developing an architecture that will allow a user to rapidly compose human performance and system models externally to the JointSAF entity simulation and have these models interact with the constructive simulation.MA&D has currently developed the architecture and many of the software modules for this effort.

These modules will allow a user to have the capability to compose a human/system performance model prior to runtime of the JointSAF simulation, and without recompiling the JointSAF code. The external model changes will be reflected within the JointSAF simulation. MA&D is also currently composing the human/system

behavior models that will interact with JointSAF. We have also implemented a subscription process that will allow the transfer of attribute ownership within JointSAF to our composable behavior simulation architecture.

2.2.3 Agent Based Measurement System for Advanced Distributed Learning Technologies

This is a Phase 1 SBIR with the Air Force Research Laboratory. MA&D's objective is to demonstrate a proposed methodology and software tool that will support performance evaluation in a Distributed Mission Training (networked simulator) environment. The proposed approach combines automated and manual strategies for capturing performance information and measuring mission essential competencies.

2.2.4 Dynamically Optimized Team Performance

A software-based system is proposed, the Dynamically Adaptive Resource Tool (DART), to help operators manage information intense monitoring systems in military-relevant operational settings. It is conceived to be an interface that uses adaptive automation to enhance information gathering operator resources that 1) displays the current effectiveness of an operator based on past or expected performance, 2) provides visual and auditory means to alert the operator to significant decreases in effectiveness, 3) adjusts information transfer rate and presentation rate based on the operator's effectiveness and 4) re-distributes task workload to other team members as needed. The objective of this approach is to integrate sophisticated workload modeling and analysis tools to help manage real-time operator performance. The DART will be able optimize the distribution of information, resources and enhance the performance of the system.

2.2.5 Tactical Warfare Instructional Support Environment

This project is to develop an integrated set of process-based performance assessment and training support tools tailored for the Sea-Based Weapons and Advanced Tactics School (SWATS) and design it so that the framework could be applied to performance assessment and training support in other military and non-military domains. The integrated tool set is referred to as the Tactical Warfare Instructional Support Environment (TACWISE).

The purpose of TACWISE is to enhance the training process and improve the user community's capability to link training to performance readiness. The tool is based on the premise that the human performance component of combat readiness, measured primarily in terms of performance outcome, can best be improved by focusing training on both the processes and outcomes of performance. By assessing the performance process in addition to performance outcomes, training interventions, such as post-exercise debrief and the design of training events can be focused on identified process weaknesses. A critical requirement for achieving this purpose is to provide support to the training community in developing and utilizing measures of performance (MOPs).

BCI, Inc. also has extensive project experience, which will aid in the development of a comprehensive readiness assessment methodology. The following projects characterize that experience:

2.2.6 Advanced Integrated Electronic Warfare System Manpower Assessment

BCI developed an innovative manpower, personnel, and training (MPT) assessment for the next generation electronic warfare system. The significance of this effort is that *this MPT is being conducted prior to design of the system.* This is an effort to allow human operator concerns to influence system design.

2.2.7 Area Air Defense Commander (AADC) Training Project Plan

BCI wrote the AADC TPP for ATRC. This document lays out the design of a course of instruction (COI) that includes: justification for the course; impact if the course is not developed; course data (title, length, Course Curriculum Model Manager (CCMM), planned student capacities, etc.); course overview (curriculum outline; milestone chart leading to first teach and instructor certification, etc.). Currently, BCI is writing the AADC prototype curriculum and coordinating the transition of that training from the lab, JHU/APL, to ATRC in Dahlgren. This effort is laying the foundation for the development of the transition from prototype training to full production system training. Additionally, BCI co-authored the AADC Capability Prototype Training CONOPS, and wrote sections to the AADC NTSP.

2.2.8 Battle Group Multi-TADIL Training course Review

BCI attended the initial teaching of the tailored Battle Group Multi-TADIL Training Course (BGMTT) for the purpose of reviewing course content and presentation. Recommendations were provided to ATRC for inclusion into subsequent tailored BGMTT courses.

2.2.9 Cooperative Engagement Capability (CEC) TACMEMO Development

BCI as part of the TACMEMO Development team provided support to PEO TSC toward the success completion of CEC OPEVAL. The development team was responsible for developing and fielding the CEC Tactical Employment Guide for OPEVAL. The employment guide provided a single, hyper-linked source for relevant guidance relating to tactical set up and employment of Battle Force Air Defense systems.

2.2.10 AEGIS Training and Readiness Center Support

BCI provides ongoing support to ATRC in the area of new curriculum development, support, and implementation of new training strategies and techniques. One such project was the implementation of the use of the ShipMate hand-held instructor device. BCI performed the training for the instructors and developed methods for the best use of the device.

Additionally, BCI provides support to ATRC, N34 – Warfare Analysis and Operations Department, providing analysis of operational capabilities and lessons learned for incorporation into the active course curricula.

3. Phase I

3.1 Phase I Team

Micro Analysis and Design, Inc. (MA&D) has a solid 14 year history of developing quality software products and tools for evaluation of human performance for both Government and commercial clients. During that time we have researched and developed innovative tools for analyzing and predicting human performance, evaluating system design and determining degradation effects on human performance due to environmental conditions. Additionally, we have been instrumental in developing useful interface guidelines for designers to enable them to take maximum advantage of technology while considering human performance capabilities.

We have unique experience in developing training, human factors and decision support tools for the Army, Air Force, Marines and Navy. In addition, we have developed tools for other military and Government organizations around the world. Our strength is in developing useful software that capitalizes on existing information bases and meets the analysts' needs to perform practical, scientifically based, defensible human performance modeling studies.

For this effort, we have teamed with Basic Commerce and Industries, Inc. (BCI). BCI has extensive operational experience with Naval war fighting systems and is currently providing top-level systems engineering support to the AEGIS program and other new acquisition programs such as LPD-17, DD21, and the Advanced Integrated Electronic Warfare System (AIEWS). In addition to their roles in support of these future combat systems, BCI provides support to the AEGIS Training and Readiness Center (ATRC) in Dahlgren, VA. The BCI team combines the experience of engineers and command level surface ship operators to provide a unique capability to determine future system performance parameters. BCI is currently supporting numerous projects involved in developing advanced concepts for combat systems, training, maintenance and command and control.

3.2 Summary of Phase I Accomplishments

3.2.1 Develop a readiness measurement methodology

The initial step for Phase I was to determine the most appropriate combination of measurement and reporting strategies to gauge crew readiness in a wide range of mission areas.

3.2.2 Demonstration of feasibility

The primary goal of Phase I was to study the feasibility of developing the mission readiness methodology into a full scale, deployable tool that could be used primarily as an automated readiness measurement system.

3.2.3 Decomposition of UNTL Tasks

A second goal of the effort was to create a hierarchy of tasks starting with Universal Naval Tasks and tracing them down to specific operator actions which could be used to evaluate performance of the Universal Naval Task.

3.2.4 Construction of initial task network model

As future weapons systems come on line and are deployed in the fleet, the operator actions required to accomplish a mission essential task, and ultimately a universal Naval task, will change rapidly. Constructing task network models of the operator actions that combine to complete a Universal Naval Task will allow for easy updates of the operator actions. In addition, the model can be set to run at an expert level and can provide an additional measure of feedback beyond just whether or not performance standards were met, but whether performance measures up to expert performance.

3.3 Phase I Tasks

3.3.1 Overview of Project Objectives

In this section we summarize the most important objectives for MIRMAID, and briefly specify how MIRMAID will achieve those goals. Details on how we intend to achieve these goals through MIRMAID will be discussed later in this report. These objectives were identified in the initial Phase I proposal and solidified in the SBIR Day kick-off meeting with our Government team members.

- Develop a readiness measurement methodology.

The first step in constructing the MIRMAID readiness assessment tool was to identify the most appropriate strategy for assessing crew readiness. The resulting combination of automated and manual performance assessment and monitoring strategies formed the foundations for all subsequent work in Phase I.

- Identify specific mission areas for targeted research in Phase I.

A main objective of the Phase I effort was to identify a small subset of mission areas that could serve as test cases for an automated readiness measurement system in Phase II of the SBIR.

- Determine the operator level tasks that must be completed to accomplish the selected mission areas.

Having identified the mission areas that would be targeted for research and exploration in Phase I it was necessary to decompose those mission areas into discrete operator actions, which could be evaluated against some objective measure of performance.

- Investigate queues in BFTT or the combat system to be used in performance assessment.

In order to evaluate the performance of an operator we must first know what tasks would be appropriate for a given scenario. To that end, MIRMAID must interact with BFTT or future embedded trainers to interpret the scenario and identify queues in a scenario, which should prompt an operator to begin a sequence of tasks.

- Create models of expert performance in the selected mission area

In addition to identifying sequences of operator actions required to perform a selected mission, a task network model was constructed of the operator performing the given tasks. This will allow the MIRMAID system to evaluate operator readiness not only against objective measures of performance or effectiveness, but also against the expert (or any other) operator.

The work completed in Phase I of the MIRMAID project has led to a functional description of a software tool that will meet the Government's objectives. We will describe the approach and functional design of the software package in greater detail later in this report.

Phase I consisted of five separate tasks. They were:

1. Develop a readiness measurement methodology
2. Identify mission areas and tasks for investigation in Phase I
3. Investigate queues/triggers in the combat system and BFTT to be used in performance assessment
4. Identify additional data collection requirements
5. Model selected mission areas
6. Develop a functional description of the MIRMAID software

3.3.1.1 Task 1: Develop a readiness assessment methodology

The first task for the project team was to develop a theoretical measurement framework on which to build the readiness assessment software. It was determined that the best readiness assessment methodology would begin with the Universal Naval Task List (UNTL) and deconstruct the constituent tasks down to the level of specific operator actions. The diagram below illustrates the hierarchy that was constructed as a basis for assessing the completion of a UNTL item.

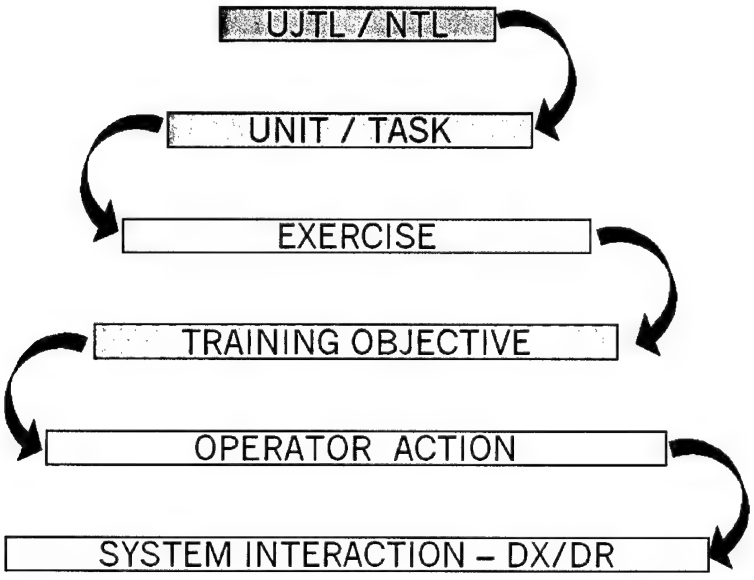


Figure 1 MIRMAID Readiness Assessment Methodology

The MIRMAID readiness assessment methodology begins with the highest, most generic level of a task required for any Navy platform, a Universal Naval Task. The UNTL represents high-level requirements that a ship must accomplish. These tasks are not platform specific, rather they represent a generic task list for many ships. At the Unit/Task level of the MIRMAID hierarchy tasks are grouped into specific warfare or

mission areas for a given ship. This set of tasks represents the specific parts of a UNTL task that a certain ship or class of ships is responsible for completing. Having identified the group of UNTL items assigned to a particular ship or class of ships the next step is to look at an exercise or event. An exercise represents a specific time when a ship accomplishes a specific set of tasks with a logical grouping. An example of an exercise would be an Air Defense Exercise (ADEX). The list of tasks or objectives that need to be accomplished for the exercise would contain all the sub-tasks from the Naval Task List (NTL) which are used to complete the tasks associated with the particular exercise. At this level, many of the tasks accomplished in completing an exercise would be re-usable and transferable to other exercises (i.e. SPY radar set up). Within any given exercise would be a set of training objectives or sub-events in the exercise process that are associated with completing a higher-level task. For example, in the ADEX mentioned above, there would be several stages such as planning/set-up, detection, track maintenance, ID management, engagement decision, engagement execution, kill assessment, and recovery. Each of these stages is a set of actions triggered by an observable event within the combat system or a timed requirement. Finally, given the stages of a particular training objective, we can trace to an individual operator which accomplishes a step in the process. Many actions and processes may be occurring at the same time at one or multiple watchstations. These actions are in turn linked to specific data extraction/data reduction (DX/DR) triggers. Triggers are required because actions occur in response to tactically relevant events in the combat system which are typically recorded through the DX/DR LAN.

In order for this readiness assessment methodology to work, triggers within the combat system must be identified for each action that flows down from the top-level requirements of the UNTL. For the Phase I effort a small set of UNTL items were decomposed down to operator actions which were then linked to specific data extraction points from the AEGIS combat system which was chosen as the target platform for the initial investigation. These specific tasks will be discussed later in more detail. Earlier efforts conducted under previous design efforts such as the SH-60 cockpit trainer, A-10 tactics trainer and Device 20A66 clearly indicate that all observable tactical decisions, tactical procedure execution and operational procedure execution is related to some observable activity which precedes it (Baumgartner, DIS Journal, 1994). In addition, the test of the automated performance measurement system explored by Baumgartner (1994) clearly demonstrated that triggers are available at a level well in excess of those needed for full and detailed information of the process as designed in the earlier effort and tested on AEGIS Baseline 5.

3.3.1.2 Task 2: Identify mission areas for Phase I investigation

Having identified the desired methodology for assessing crew readiness, the next task was to identify a small subset of mission areas and tasks for demonstration of the chosen assessment methods. Identifying target mission areas and tasks for the Phase I effort began with the Universal Naval Task List.

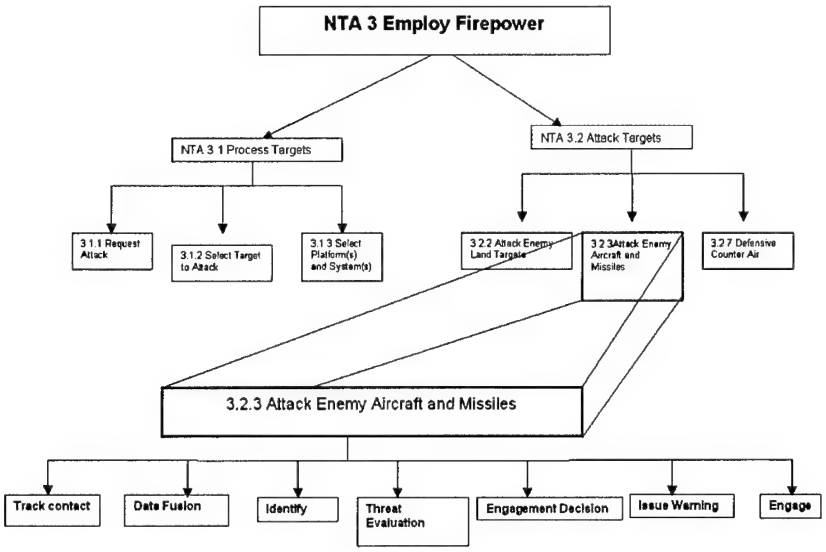


Figure 2 NTA 3 'Employ Firepower' Deconstruction

3.3.1.3 Task 3: Investigation of queues and triggers for automated data collection

The cornerstone of the MIRMAID system will be an automated performance assessment and monitoring capability. Automated performance measurement is made possible through deconstructing UNTL items down to individual operator level tasks which can be tracked through triggers in the DX/DR LAN or the BFTT interface with this network. In order to accomplish this the MIRMAID system must be equipped to passively tap either the DX/DR LAN or BFTT to harvest the requisite data. Figure 3 below shows the generic system architecture of a Navy vessels today with the addition of the MIRMAID readiness assessment software package. As the diagram shows there are two potential places where MIRMAID may be able to passively tap into the existing ship systems architecture to harvest performance data. Each of the options has its strengths and weaknesses and the ultimate decision concerning where MIRMAID collects data should be made in concert with our government partners at the outset of Phase II.

The first option for tying the MIRMAID system into the ship's systems is to passively tap the communication link between the host combat system and the DX/DR LAN. Taking data as it passes from the combat system to the DX/DR LAN is advantageous for many reasons. First, placing the passive tap in this position ties MIRMAID into a communication line with an existing two-way interface. That is, not only can MIRMAID take data as it passes from the combat system to the DX/DR LAN, but also in the future MIRMAID will be able to serve as the foundation for intelligent systems that can aide operators in performing their jobs through dynamic feedback. In addition to the advantages of the two-way communications link directly to the host combat system, using the passive tap between the combat system and the DX/DR LAN has the advantage of being independent of BFTT. This may be particularly important during the early deployment of MIRMAID when the primary goal of the system will be to collect

data that can be used later for performance evaluation. Tapping into the system directly will allow MIRMAID to collect data during battle conditions rather than only in a training context if MIRMAID accesses data through BFTT.

Despite the aforementioned advantages for accessing data through the link between the host combat system and the DX/DR there are drawbacks to this approach as well. The primary drawback of tapping directly into the link between DX/DR and the host combat system is the variety of host systems MIRMAID will work with. Extracting MIRMAID data from the link between combat system and DX/DR would force MIRMAID to convert the data from multiple combat systems into a uniform format for processing additional code development and research to the overall project. This problem is remedied by harvesting performance data directly from BFTT, as the data will have a uniform format that is readily available in Appendix K of the BFTT manual.

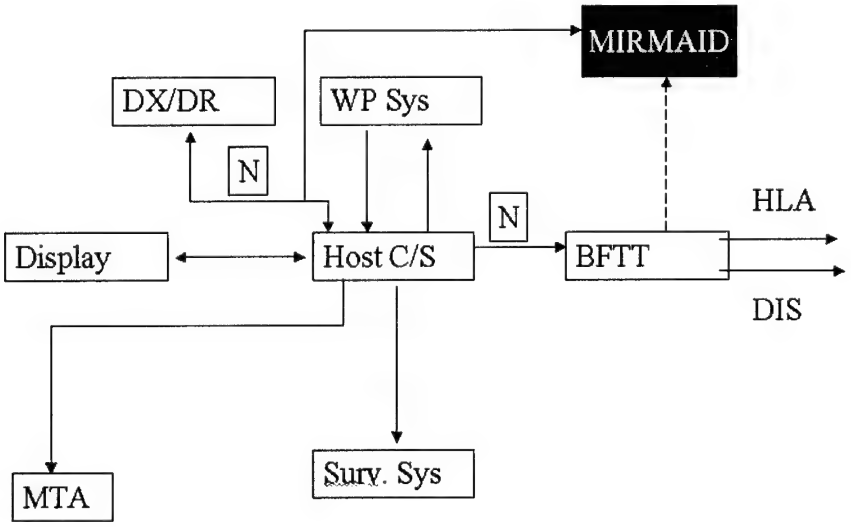


Figure 3 Generic Naval System Architecture and MIRMAID

As illustrated in Figure 2, specific operator tasks can be identified that when completed can lead to the conclusion that some higher level Navy task has been accomplished. The challenge here is to map those operator actions to data extraction points that can provide relevant performance data. For Phase I, the data extraction points from the AEGIS combat system served as the test set which allowed for identification of specific extraction points to measure the operator actions from the chosen mission areas. For example, as seen above the breakdown of UNTL item NTA 3 identifies the operator tasks of: track contact, data fusion, identify, threat evaluation, engagement decision, issue warning, and engage. For many of these tasks extraction points were identified from the AEGIS combat system, which could indicate either the start or completion or both of the operator tasks. Table 1 below contains a sample list of data extraction points taken from the *AEGIS Baseline 6A0.1.1 Data Extraction and Reduction Guidelines (TW271-AE-GYD-260/MK7)*. Each extraction point (EP) listed below indicates data that the AEGIS combat system sends out to the DX/DR LAN that

contains information useful in monitoring the performance of operators in CIC. The content of these data extraction points will be collected by MIRMAID and can then be used to evaluate performance against pre-defined measures of performance and effectiveness, or stored for later use in analyzing trends in task performance.

Tasks	Track Contact	Data Fusion	Identify	Threat Evaluation	Engagement Decision	Issue Warning	Engage
Data Points	EP 46		EP 58 EP 99 EP 104 EP 105 EP 127			EP 326 EP 327 EP 350 EP 388	EP 32 EP 201 EP 207 EP 221 EP 227 EP 228 EP 232 EP 288

Table 1 Sample AEGIS Extraction Points for NTA 3

3.3.1.4 Task 4: Identify additional data collection requirements

While MIRMAID relies heavily on data that is currently collected by BFTT through the data extraction and reduction capabilities within each host combat system, there are performance parameters that are either outside current data collection or may not be appropriately measured through the automated system. Through the investigation of the prototype UNTL items, some additional data collection requirements were identified. In order to accomplish a robust readiness assessment methodology current data collection will need to be augmented with additional automated performance measurement capability and possibly some manual data collection by an individual supervising the training exercise.

The primary use for manual data collection would be the evaluation of verbal communications that cannot be captured in an automated fashion. Engagement orders, track identifications, and emitter to track correlations passed over verbal communications links, are essential to team readiness, but have automated data available for their assessment. In order to capture this valuable aspect of team performance MIRMAID will include in its overall readiness reporting and assessment capability the option for a training supervisor to rate certain aspects of performance through a hand-held computer. The specifics of the manual data collection portion of the total MIRMAID system will be detailed in the functional description of MIRMAID provided in Section 3.4 of this document.

In addition to identifying performance data that will require manual evaluation the investigation performed for Phase I identified certain parameters which may be need to be added to current automated data collection.

The first area identified as a possible opportunity to expand current data collection is the recording of activity on voice networks. One of the main workload drivers for manual evaluation in MIRMAID will be the necessity for evaluation of voice traffic in the CIC. The idea behind adding voice recording to the data collection capability would be a vision for future technology to possibly take over evaluation of voice communications. Currently, voice recognition and interpretation software provides

only limited capability to understand and evaluate oral communication. However, as technology in this area continues to move forward we may be able to off load some of the manual data collection to an automated system. Incorporating the data collection capability early on in the development of MIRMAID will allow for testing of emergent voice recognition technologies without the cost of rebuilding MIRMAID or the system from which it harvests data.

A second major area of need in the automated data collection process is that of equipment fault or failure. As automation increases and crewmembers become more and more reliant on computer systems the proper functioning of deployed systems becomes increasingly more important. Not only is the ability to function in the absence of mission critical equipment a large concern, but identifying the specific areas of performance that are effected by loss of equipment is paramount. By collecting equipment fault and failure data and feeding it to MIRMAID we can begin to identify performance trends that can be related directly to certain pieces of equipment. Additionally, enabling MIRMAID to show training supervisors when equipment failed and the resulting performance impacts will help the instructor to identify and improve performance in areas that are directly affected by equipment failure.

It is important to note that because the initial investigation into developing the MIRMAID system focused on a limited set of UNTL items there may be additional data collection requirements that emerge as the coverage of MIRMAID expands. However, the investigation and resulting performance assessment methodology from Phase I indicate that current data collection provides a substantial foundation for building MIRMAID. The research into the data extraction and reduction capabilities currently available from the AEGIS combat system and BFTT suggest that any significant extension of data collection would likely be unnecessary.

3.3.1.5 Task 5: Model Selected Mission Areas

A unique feature of MIRMAID is the inclusion of expert models to provide performance-monitoring, feedback and identify where to focus training resources. Previous training and performance assessment efforts (e.g. AET) have included models of expert performance for the purpose of comparing operator performance to desired expert levels and proven that such an approach is valuable in Navy training. However, unlike the cognitive modeling approach of the AET program, MIRMAID will utilize behavior based expert models in performance evaluation.

The expert models for the MIRMAID performance assessment system will be built using the Micro Saint simulation engine. Micro Saint is a discrete event, task network-modeling tool designed by Micro Analysis and Design, Inc. Models are created through defining a chain of tasks that are executed in sequence and triggered by some event. In MIRMAID, when a data point is passed from the combat system that indicates a certain event is happening the Micro Saint model will be triggered and will simulate the actions the operator should take in reaction to the scenario. Using a behaviorally based model has many advantages. First, behavior based models remove the need to make inferences about cognitive processes that accompany overt behaviors. The Micro Saint models used for MIRMAID will focus on observable behaviors that can have an identifiable beginning and ending point and that can be measured in some meaningful way to aid in performance evaluation. A second

advantage of the task network modeling approach chosen for MIRMAID is that the models are easily updated. Over time, as new systems are brought on line and performance standards are determined, the behaviorally based models could be easily updated to reflect new tasks or task times without a complex rebuilding process. A final advantage of the expert models utilized in MIRMAID is the relatively short development time line. An initial model for a selected mission area can be constructed in a matter of weeks and once constructed the models can be updated in days. In contrast to other modeling methods Micro Saint models can be developed and maintained in an efficient and effective manner to ensure that they accurately reflect the latest developments in the fleet.

For Phase I of the MIRMAID project an expert model was constructed to support performance assessment of NTA 3.2.3 Attack Enemy Aircraft and Missiles. The model was built using inputs from subject matter experts who assisted in identifying the task progression associated with completion of this task and nominal times for each individual task. Figure 4 below is the preliminary model constructed during Phase I. With MIRMAID running a training scenario the models will process tracks with ground truth information and evaluate the timing and decisions of the operator. However, if MIRMAID is running in real combat situations for data collection and performance evaluation under realistic conditions, the model is designed to run at speeds faster than real time so that as additional data is collected on a track the model can update the track information and be re-run to reflect more current conditions.

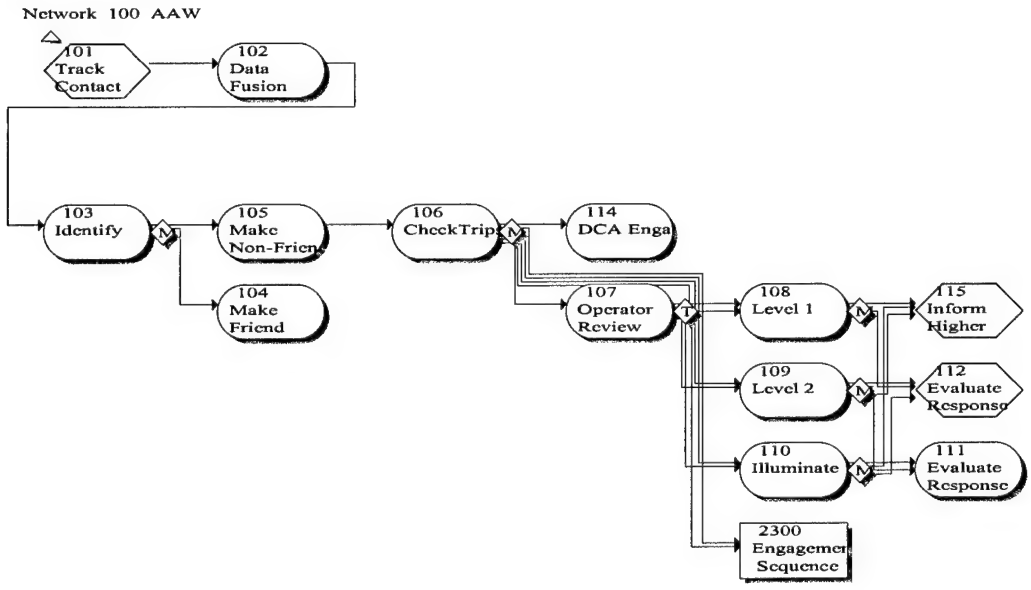


Figure 4 Expert Model for NTA 3.2.3

The MIRMAID system will be designed with the expert model portion of the system as an adjunct to the overall readiness assessment methodology. Such a design will allow for the system to function in areas where there is insufficient data to construct expert models and will afford the training supervisor the option of including expert model feedback as part of their performance evaluation.

3.3.1.6 Task 6: Develop functional description of MIRMAID product- Phase I product

The final task performed under the Phase I contract was to develop a fully functional description of the MIRMAID tool and its capabilities. Along with this functional description we identify additional key features of the software, such as outputs of the system and a potential graphical user interface for the final software. This functional software description is the primary deliverable for the Phase I effort, it will serve as a major component of the Phase II proposal which will outline the work plan that will take MIRMAID from the conceptual level into a deployable tool.

4.0 MIRMAID Functional Description

The overall design strategy for MIRMAID will focus on five main items:

- Use of existing data collection – the main goal of the MIRMAID readiness assessment methodology is to be able to leverage existing performance data collection. Identifying the most appropriate location for MIRMAID to passively tap the existing ship's systems architecture to harvest data that is already being collected and passed through the ship's LAN is paramount to the ease of deployment sought with MIRMAID.
- Integration of manual data collection – While MIRMAID's main purpose is to provide automated performance measurement capability, the contractor team fully realizes that certain aspects of individual and team performance cannot be gauged through automation. With that in mind, MIRMAID will provide an integrated manual data collection option that will allow training supervisors to rate performance via a hand-held computer and incorporate these ratings into overall readiness reporting.
- Use of existing readiness reporting formats – A third goal of the MIRMAID system is to allow easy transition into deployment. Resistance to change is often a major factor in deploying new systems. In order to lessen the resistance to use of the MIRMAID system the readiness reports will be designed to closely emulate current readiness reports.
- Data collection capability – Given that performance standards may not be available to all areas MIRMAID will need to offer a data collection system which can be used to gather performance data and assist in the determination of ideal and acceptable levels of performance. This would require identification of key performance parameters that could be monitored, recorded and analyzed over time.
- User defined performance standards – In addition to collecting actual performance data to establish threshold measures, MIRMAID will allow the user to define objectives. This capability will allow training supervisors to evaluate team and individual performance at varying levels and enable them to alter performance standards based on their current stage of the inter-deployment training cycle.

The most efficient discussion of the functional capability of MIRMAID can be accomplished through discussion of a series of graphics designed to represent the functional flow of MIRMAID and possible user interfaces. In the following

pages we will present a series of diagrams and pictures of potential screen shots and discuss each of them in order to detail the functionality of MIRMAID.

4.1 MIRMAID Functional Flow

Figure 5 below represents a conceptual look at how a user would proceed through the MIRMAID tool. The top block in the figure represents the pre-defined training scenario on which the performance will be evaluated. MIRMAID will be designed to run in conjunction with a variety of currently deployed training systems. When a scenario is loaded into a training system MIRMAID will accept the training scenario and when possible present the user with a list of UNTL items that may be exercised in the given scenario.

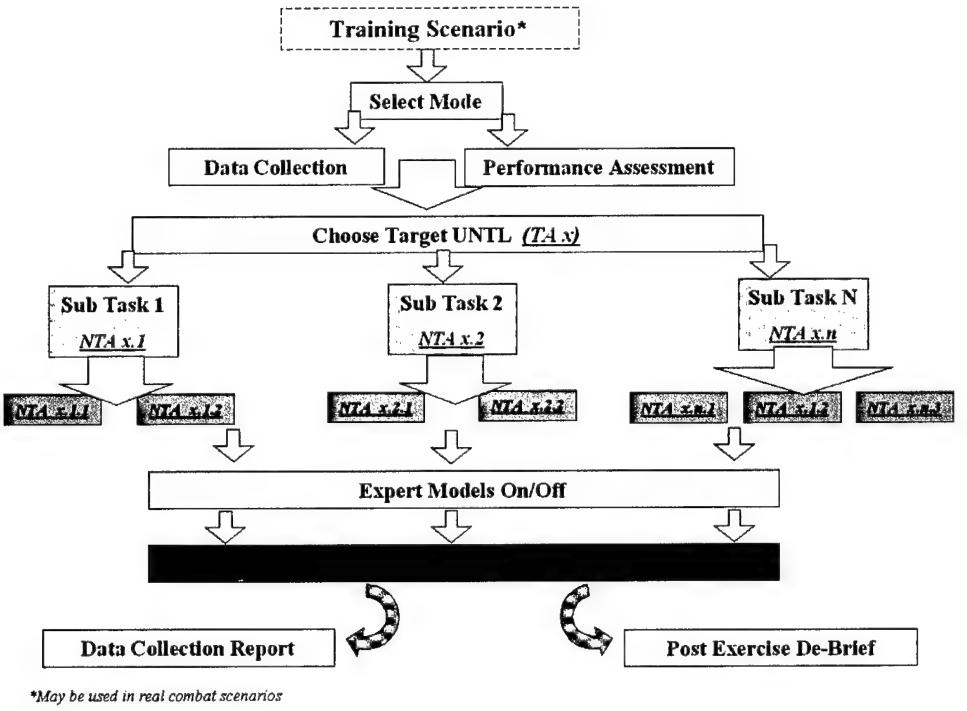


Figure 5 MIRMAID Functional Flow

The first choice a user will make when entering MIRMAID will be whether the system will operate in "Data Collection" or "Performance Assessment" mode. The primary effects of the user's choice of mode will be seen in the output from MIRMAID. Final MIRMAID outputs will be discussed in greater detail later in this section, but in brief, Data Collection Outputs will consist of statistics describing crew performance on selected measures, their related sub-tasks and the higher-level UNTL items. Outputs generated while in "Performance Assessment" mode will be full debrief reports to provide instructors with evaluative information on the training exercise and facilitate debrief of the exercise with the participants.

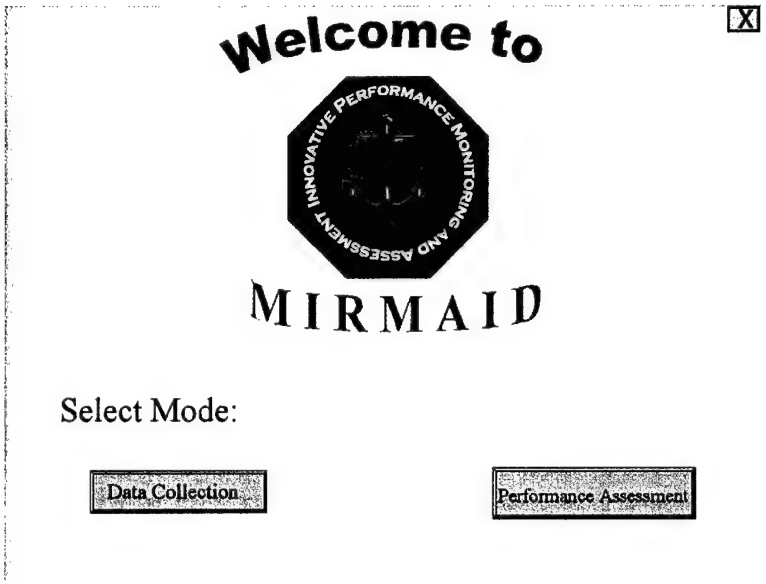


Figure 6 MIRMAID Welcome Screen

Having selected the mode of operation for the MIRMAID system, users will then select specific UNTL items they wish to assess during the given training exercise. Ideally, UNTL items will be linked with specific training scenarios so that after selecting the system mode, users will be presented with a list of candidate UNTL items that are eligible for evaluation given a particular assessment. While operating in Data Collection Mode, MIRMAID will give the users a "Select All" option to collect performance data on all relevant UNTL items. In addition to allowing users to designate UNTL tasks for evaluation during a particular exercise, users will be able to designate data collection based on a more traditional warfare area breakdown (i.e. AAW, ASW, SUW, etc). Regardless of the method used to designate performance measures for evaluation in MIRMAID, final reports will reflect these measures in relation to the UNTL tasks which they satisfy. In addition to specifying data collection based on UNTL, users will be able to follow the pre-established hierarchies from UNTL down to specific operator tasks and define data collection at the lowest level of operator action.

After specifying the areas for which data will be collected in a given exercise, users will be given the option of defining specific measures of performance or effectiveness for the crew under the current scenario. Early on in the deployment of MIRMAID, the only available performance standards will be those entered by the user. However, as MIRMAID becomes more and more widely used, the embedded data collection function will allow users to utilize past performance data to set MOP/MOE's. Once the user has defined the MOP/MOE's the training scenario can be executed, with the additional level of expert model feedback enabled if so desired, and outputs examined.

4.2 Data Collection Mode

As described above, users may wish to operate the MIRMAID system in "Data Collection" mode. When operating in data collection mode, all functions of MIRMAID available to the user in "Performance Assessment" mode will be available. The primary

difference between the two operational modes of the MIRMAID system can be found in the outputs associated with each mode.

In data collection mode, users will not define specific MOP/MOE standards for evaluation in the final report. Rather, after choosing specific areas in which to monitor performance, the training scenario will execute and performance data will be stored in the MIRMAID performance database. After execution of the training exercise, users will be able to view raw performance data for each performance domain previously selected and will also be able to produce statistical reports of past performance on the selected MOP/MOE.

It is important to note that while MIRMAID is operating in data collection mode, the training supervisor will have the option of completing the manual data entry forms and adding those evaluations to the MIRMAID database. However, in data collection mode, the user can choose not to complete the manual data collection process.

4.3 Performance Assessment Mode

The main purpose of MIRMAID is to provide an automated performance assessment capability that can work in concert with a variety of training systems currently under deployment. Placing MIRMAID into the "Performance Assessment" mode will satisfy this objective.

Upon entering the performance assessment mode in MIRMAID users will be prompted to select whether they want to define their MOP/MOE's by selecting UNTL tasks or by warfare area. If the users choose to select target performance areas by warfare area they will be presented with a list of specific MOP/MOE's that are related to the chosen warfare area. Then the user will select a performance measure and be given the opportunity to define a standard for that measure. As MIRMAID matures and more data is available, default values for each MOP/MOE will be provided to the user to either accept or change for the current exercise.

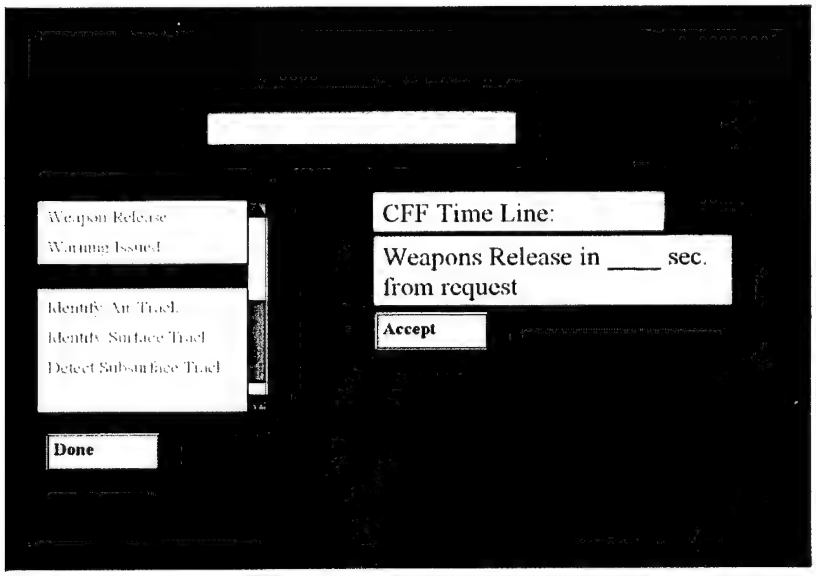


Figure 7 Sample Define MOP by Warfare Area

While users will have the ability to select target performance areas by warfare area, MIRMAID will encourage the user to utilize the pre-defined UNTL to operator action hierarchies to select and define areas that they wish to assess. The idea behind steering training supervisors toward the UTNL method of defining performance objectives is to begin to create a universal readiness assessment methodology that can be used aboard any Navy vessel. In addition, by defining performance objectives beginning with UNTL tasks, users will gain an understanding of how specific operator actions relate to high level operation requirements.

If the user chooses to work through the UNTL hierarchies, they will be presented with a list of UNTL tasks based on the currently loaded training scenario. After selecting the UNTL tasks they wish to examine, the user will be prompted to select specific subtasks that relate to the higher-level tasks. Users will proceed down the UNTL hierarchies until they reach the lowest level of operator action. Then they will be able to define specific standards for performance on tasks of their choosing.

After defining the specific areas of performance for evaluation in the current training exercise users will be asked whether they wish to include expert model feedback in the current performance analysis. The user may choose to either allow the models to execute in the background of MIRMAID or they may opt to eliminate the comparison of crew performance to the expert models for the current analysis.

4.4 Manual Data Collection

Along with the automated data collection and performance assessment functionality of MIRMAID, users will have the option of completing manual data collection forms to augment the automated system. In the process of decomposing UNTL tasks certain aspects of performance were identified that cannot be captured through automated data collection. In order to completely assess performance MIRMAID will allow the user to measure these aspects of performance through manual evaluation through a hand-held computing device. The primary function of MIRMAID's manual data collection segment will be to allow the training supervisor to rate team performance skills and oral communication skills. The combat information center is characterized by a great amount of team interaction and verbal communication. With no automated assessment currently available, these vital aspects of performance need to be captured through manual evaluation. However, manual performance assessment can become a cumbersome, labor-intensive process that requires multiple raters for each training exercise. To combat that tendency, the manual data collection in MIRMAID will be built with the following characteristics:

- Manual data collection will provide an additional level of performance feedback. Training supervisors will be able to operate MIRMAID in a fully automated measurement mode eliminating the requirement for close, direct supervision of all training exercises.
- Manual data collection will be available in both performance assessment and data collection mode. While MIRMAID is operating in the data collection mode, training supervisors may need to focus less on observing crew performance and preparing for an exercise debrief and thus better able to collect manual data. In addition, collecting

manual performance data during a data collection exercise will allow users to establish standards of performance on manually rated competencies as the MIRMAID system gathers standards data for other aspects of performance.

- Manual data collection will be user friendly and completed on a hand-held computing device. Users will be presented with categories of team performance variables and communication types. Upon selecting the category to be rated MIRMAID will present the user with a behaviorally anchored rating scale (see Figure 8). These scales will allow users to make quick evaluations of observed behavior and attempt to standardize ratings across evaluators.

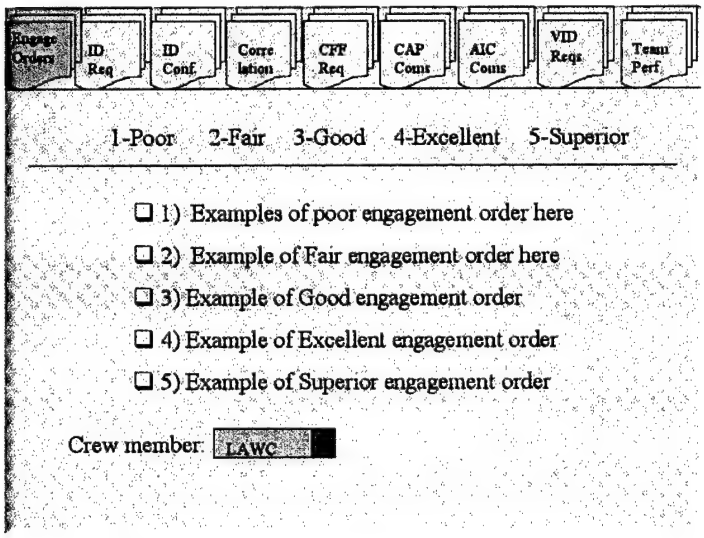


Figure 8 MIRMAID Manual Data Collection

Meeting the goals set out for the manual data collection functionality in MIRMAID will provide a powerful method for evaluating and capturing robust performance data on all aspects of crew performance. Yet, allowing users the option of including this piece of performance assessment provides flexibility to the overall readiness assessment process.

4.5 MIRMAID Output

The goal of the MIRMAID system is to add value to the performance measurement process. Regardless of what innovations are created in the MIRMAID system, the value of these innovations can be measured through evaluating the usefulness of the outputs MIRMAID provides to its users. In an effort to make transition to the MIRMAID system as easy as possible, we investigated current training reports to evaluate their usefulness for recording and reporting the outputs of MIRMAID. The first output from MIRMAID that a user will see will mirror those of current training reports utilized in the fleet today. This report will provide a quick look at the specific training exercise completed, an overall training rating and a brief explanation of the rating that outlines areas where team performance did not meet expected levels. Preserving the

format of this initial report will allow instructors to quickly print or forward the outcomes of a training exercise.

In addition to the traditional training report format presented as page one of MIRMAID output, MIRMAID will provide more detailed feedback to the user regarding team performance and assist in exercise debriefing through an electronic interface. In keeping with the MIRMAID principle of hierarchical breakdowns of UNTL tasks, the output reports will be based on a similar principle. MIRMAID will present the user with a report of the UNTL tasks that have been assessed and trace performance from the lowest level of operator action back up through the hierarchy and report performance up to the highest level appropriate for the given exercise.

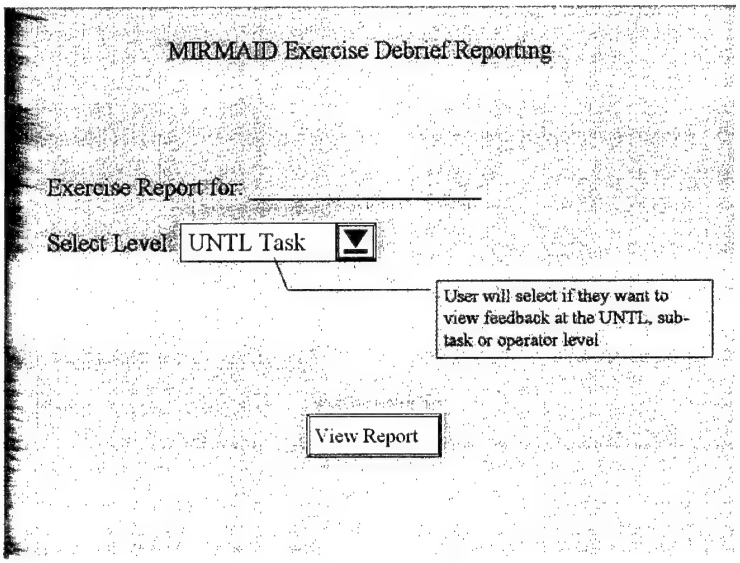


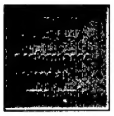
Figure 9 MIRMAID Output Entry Screen

Users will be able to select the initial level of feedback displayed in the output report and then navigate to additional levels through hyperlinks in the report format. In addition to displaying performance scores, MIRMAID will allow the user to access statistical data that compares current performance to past performance on the same UNTL task as exhibited in other scenarios. MIRMAID will also provide the user with information regarding how often each specific task and sub-task was assessed in a given scenario and how often performance met expected levels, along with the overall performance assessment represented in the sample screen below.

Exercise Report for: _____

Select Task: NTA 3.1

Sub-Task	Standard for Current Exercise	Actual Performance
NTA 3.1.2 Attack Land Target		
NTA 3.1.3 Attack Surface Targets		
NTA 3.1.4 Destroy Enemy Aircraft		



Instructor Comments: This space will show any written comments made by the users while conducting manual evaluation

Figure 10 Sample Output Screen

The strength of the MIRMAID output format as exhibited here is the flexibility it affords the user. Users can decide at what level of detail they wish to receive feedback, and have the option of exploring multiple aspects and levels of detail in a single report. However, if the user only desires a quick look at team performance, the initial standard training report format provided by MIRMAID will allow the user this convenience.

4.6 Summary

The preliminary design concept described in this section reflects the initial exploration into automated performance measurement conducted during the Phase I research effort. The methodology chosen provides a strong basis for assessing crew readiness in a highly automated fashion with limited need to additional data collection and relatively little intrusion into system operations. The concepts presented here demonstrate that MIRMAID can accomplish the goals set forth for this effort. MIRMAID provides a comprehensive readiness assessment methodology while remaining flexible and adaptable to new systems.

As future work on MIRMAID proceeds, GUI development and user needs will be further explored to ensure that MIRMAID provides the best and easiest to use performance measurement system with minimal impact to operational crewing aboard Navy vessels.

Appendix A – Additional UNTL Decompositions

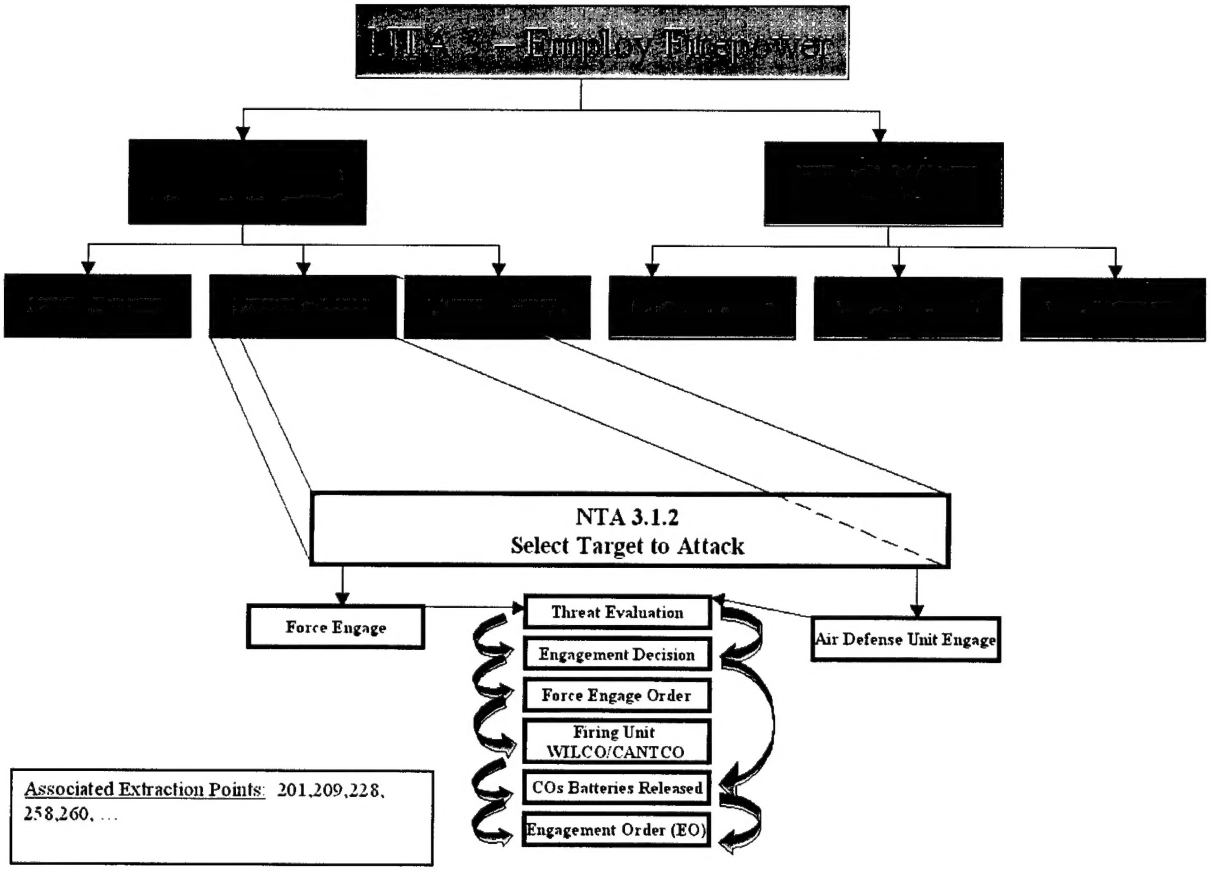


Figure 11 Example Breakdown of NTA 3.1.2

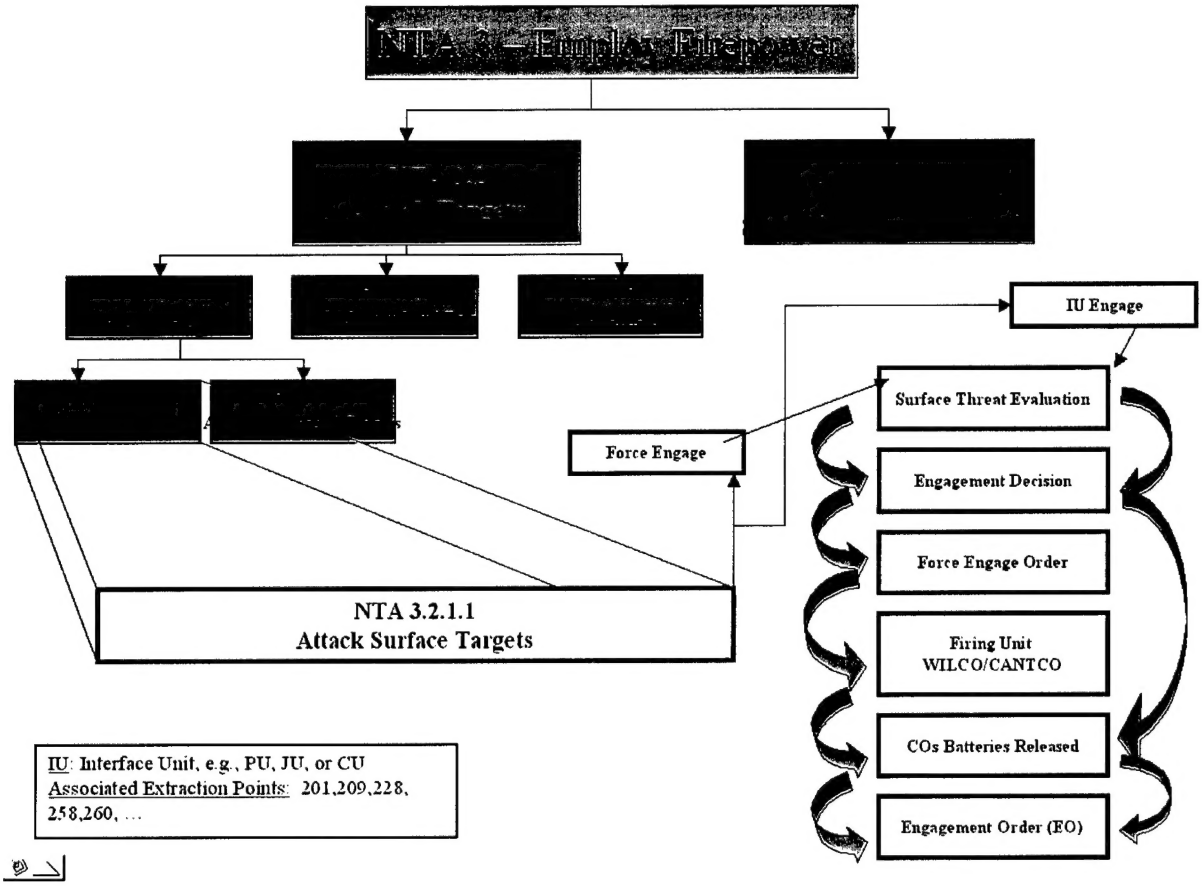


Figure 12 Sample Breakdown of NTA 3.2.1.1

Appendix B MIRMAID TRAINING REPORT

UNCLASSIFIED JIEOH 9000
SET USER FORMATS (4.0) STATUS: AGREED
IMPL DATE: 31 MARCH 1999 DATE: 12DEC1991

MEASURED CATEGORY LEVEL FOR TRAINING

TRAINING/_____/...../
1 TRAINING CATEGORY LEVEL CODE 2 CODED REASON TRAINING NOT C-1
M [1 N] C [3 AN]

EXAMPLE: TRAINING/TRRAT:2/TRRES:T22//

FIELD FILLER NOTES

NO	OCC	DESIGNATOR	EXPLANATION / ALLOWED FORMATS / EXAMPLES	STRUCTURE	FLD-DESC
1	M	TRAINING CATEGORY LEVEL CODE	ENTER THE EXTENT, BASED ON ACTUAL MEASUREMENT, THAT A UNIT HAS ATTAINED THE PRESCRIBED LEVEL OF TRAINING TO UNDERTAKE THE WARTIME MISSION FOR WHICH IT IS ORGANIZED AND DESIGNED. USE TABLE 2115A FOR ALLOWABLE ENTRIES EXAMPLE: /TRRAT:2	1 N	TRRAT
2	C	CODED REASON TRAINING NOT C-1	NOTE 1: FIELD 2 IS MANDATORY IF FIELD 1 EQUALS "2", "3", OR "4". ENTER A CODE THAT REPRESENTS THE MOST SIGNIFICANT REASON WHY THE MEASURED TRAINING RESOURCE AREA IS NOT CURRENTLY AT CATEGORY LEVEL CODE 1. USE TABLE 2355B FOR ALLOWABLE ENTRIES EXAMPLE: /TRRES:T22	3 AN	TRRES

REVISION DATE: 18JUL1997 TRAINING-1
UNCLASSIFIED